DOCUMENT RESUME

ED 171 553 SE 027 697

AUTHOR Davis, Robert B.; McKnight, Curtis C.
TITLE Towards Eliminating "Black Boxes"; A New Look at Good vs. Poor Mathematics Students.
PUB DATE Apr 79
NOTE 20p.; Paper presented at the annual meeting of the American Educational Research Association (San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Artificial Intelligence; *Cognitive Ability; *Educational Research; Elementary Secondary Education; *Information Processing; Interviews; *Learning Difficulties; *Low Achievers; *Mathematics Education; Models; Postsecondary Education; Problem Solving; Research Methodology

ABSTRACT

This study, based on the analysis of extensive interview data using concepts and models drawn from cognitive studies of human behavior and from artificial intelligence studies of computer information processing, highlights several information-processing differences between strong and weak students. Data were collected from 300 students in grades 3 through 10 (but mainly in grades 7, 8, and 9) and a small number of adult students in a local community college. The clinical interview technique was used, during which subjects attempted to solve problems in arithmetic, algebra and geometry. Extensive analysis of these observational data were used to identify patterns of information-processing acts providing useful explanations of the differences between good and poor mathematics students. (HM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED171553

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Robert B. Davis

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TOWARDS ELIMINATING "BLACK BOXES": A NEW LOOK AT GOOD VS. POOR MATHEMATICS STUDENTS[1]

Robert B. Davis and Curtis C. McKnight
University of Illinois, Urbana/Champaign

Mathematics educators have for several decades suffered under the frustration of dealing primarily with inputs of various sorts - instructional programs, amount of time assigned to instruction, etc. - and with outputs of various sorts - for example, selections on multiple choice test items - but with little or no theory of how inputs are related to outputs. The mind of the student was a "black box," whose internal workings were entirely unknown and unknowable. The history of other scientific studies very strongly suggests that an input-output approach is a no-win situation. Without process models very little progress can be expected.[2] Can mathematics education hope to get beyond the "black box" approach?

[1]The research reported here was supported by NIE Grant No. HEW/NIEG76 0085 and NSF Grant No. NSF SED 77 18047.

[2]Any number of cases could be cited; a particularly dramatic one deals with recent studies of diabetes, specifically with the question of how important it is to control blood sugar levels. The real cost of diabetes is particularly in blindness, heart disease, kidney damage, and other sequelae. If blood sugar level is carefully controlled, can these sequelae be avoided? Twenty years of statistical studies, re-analyses of data, etc., have failed to yield an answer. Nor is there much hope for future statistical studies: "Most investigators agree that a censensus on how to treat adult-onset diabetics, if it comes at all, will come not from resolving the UGDP controversy [a large scale statistical study], but from basic clinical research." [Kolata, 1979] Such studies deal with the way glucose attaches to hemoglobin, how glucose enters nerve and lens cells in the absence of insulin, how it cannot enter fat and muscle cells unless insulin is present, and so on. In short, it deals, not with "inputs" of insulin or degrees of control, and "outputs" (or "dependent variables") in terms of cataracts, blindness, etc., but rather with the internal mechanisms that show, in detail, the processes by which the inputs produce the outputs. [Kolata, op. cit.]

SE 027 697

We would argue that the emerging methodologies of cognitive science offer real promise for carrying us beyond the black-box stage. Both cognitive studies of human behavior, and artificial intelligence studies of computer information processing, are coming to provide a theoretical framework within which one can consider, in some detail, how students think about mathematics.

Data collection methodologies have been extended, as well, so that observation and description of what students do, and the use of task-based "clinical" interviews (a la Erlwanger and others), are now increasingly recognized as essential methods of data collection.[3]

The research described here attempts to answer the question: what do process observations and cognitive analyses tell us about the difference between the way successful students think about mathematics, as against the way unsuccessful students do.

Theoretical Framework. The basic theoretical entities are adapted from artificial intelligence, and include "procedures" (or "algorithms"), "schemata," "cues," "sub-procedures," etc. Particular use is made of a formulation by Rumelhart and Ortony, which has similarities to conceptualizations by Minsky, Papert, Schank, Abelson, Norman and others.

Methods. The main data-collection method is the "clinical interview" during which subjects attempt to solve problems in arithmetic, algebra, and geometry. Because of the interactive nature of such interviews, the interviewer can try to resolve ambiguities by asking the subject to use the same method on some new problem, devised to reveal more sharply the precise method which the subject is using. Interviews are recorded by audiotaping, by observer notes,

[3] It is interesting to note that even one of the creators of the psychometric movement, none other than Binet himself, used "clinical observation" as a necessary method to find out what students were actually doing. (cf. Binet, 1903). See also Burton and Brown, 1978.

and by the preservation of student written work, and are supplemented by classroom observations.

Extensive analysis of this observational data is used to identify patterns of information-processing acts, which are then interpreted in A.I. terms to form a "first-approximation" theory. Observations on highly successful students are compared to observational data on unsuccessful students, in order to identify differences.

Data Source. The subjects were 300 students in grades 3 through 10 (but mainly in grades 7,8, and 9), some of whom have been studied for as long as four consecutive years, and a small number of adult students in a local community college.

1. A Few Key Concepts. Before considering differences between stronger and weaker students, we list a few of the key hypothetical components in our conceptualization of information processing, especially the VMS sequence, integrated sequence, frame (or schema), and meta-language.

   a. The Procedure. A procedure is a definite sequence of steps, aimed at achieving some quite specific result. (Think of it as a human information-processing analogue of a program in a programmable hand-held calculator.) A procedure can relate to other procedures in several ways, the most important being that Procedure A can call upon Procedure B, in order to accomplish part of the assignment of Procedure A. In such a case Procedure B is said to be a sub-procedure of Procedure A, and A is said to be a super-procedure. A well-known result from Erlwanger's studies of 4th, 5th, and 6th grade arithmetic students is that their arithmetic errors are nearly always located in super-procedures; the various sub-procedures, in the main, tend to be reasonably reliable.

b. The VMS Sequence. Many mathematical tasks have a certain quasi-periodic structure: one sees something, which leads to the retrieval and execution of some procedure; the execution of this procedure yields a modified visual input, which leads to the retrieval and execution of the next segment of procedures, and so on. One thus has a sequence of visual input, procedure execution, new visual input, procedure execution,...Such a pattern is more or less clearly recognizable in, say, a long division problem in arithmetic, or a "long" multiplication problem, or in the usual written algorithms for multiplying or dividing polynomials. [It is also familiar outside of mathematics, e.g., in driving somewhere if the route is complicated and is relatively unfamiliar. One hopes to recognize a landmark - oh, yes, at that school building one turns right - then to drive along until another key landmark triggers the next key decision, and so on. One could not describe the entire route merely from memory without the prompts given by key visual inputs along the route.] Sequences of this type have come to be called visually-moderated sequences ("VMS" sequences).

c. The Integrated Sequence. Successful mathematics students, through practice, are able to develop many VMS sequences into holistic entities which they can contemplate without going through the entire sequence, step by step. (Similarly, our driver becomes able, with practice, to describe the entire route in detail - and hence can instruct others in exactly how to follow the route - without the need to physically re-trace the route in order to prompt his memory.)
Sequences which have reached this state are called integrated

sequences. [For most present readers, long division is probably an example of an integrated sequence. One can "see the whole process in the mind's eye" - at least in general - without actually writing it down on paper. The specific numerals probably must be suppressed, but the general pattern can be visualized clearly.]

d. Frames, Scripts, or Schemata. Although differences are distinguishable between the precise usages of Minsky, Schank, Rumelhart and Ortony, etc., no distinction will be made here, either in name ("frame," "script," or "schema") or in detail description. What is common to all three approaches is far more important than their differences.

Several considerations prompted the introduction of frames. For example, consider the behavior related to "understanding" a written passage, such as:

> It was Paul's birthday. Jane and Alex went to buy presents. "Oh look," Jane said, "I'll get him a kite." "He already has one," Alex said. "He'll make you take it back."

Most U.S. adults, having read that passage, can answer questions such as:

Q1: Why are Jane and Alex buying presents?
A1: Because it's Paul's birthday.
Q2: For whom are the presents intended?
A2: Paul.
Q3: Where did Jane and Alex go?
A3: Presumably to a store that sells kites and other things children like.
Q4: The next to the last word in the paragraph - "it" - to what does that refer?
A4: To the kite that Jane was proposing to buy.

Having answered such questions, most humans are surprised to find that not one of those answers is given explicitly in the written passage. The answers have come from the reader's

general knowledge of birthdays, presents, stores, etc. In Minsky's language, the reader "retrieved a birthday frame." Frames play an important role, because they embody a representation of a small part of the structure of the world. They are hypothesized to involve variables, with default procedures for assigning values to variables when input data fails to do so. [Asked to repeat the story a week or two later, most adults will unwittingly add to it, by inserting information from the "birthday frames" that was not included in the original story.]

There is a growing body of evidence suggesting the operation of some mechanism generally similar to "frames." The response to natural-language text is among the strongest. In addition to text studies like "birthday," consider also the well-known phenomenon of person A claiming that B told him fact X, when B denies having done so.

e. Meta-Language. Confronted by a problem, a student may need to retrieve one or more frames, and sequence them properly. Possession of appropriate descriptors (a "meta-language") for these frames can help. Consider the problem:

> Find the distance from the origin O (0,0,0) to the plane through the points A(5,0,0), B(0,3,3) and C(0,0,2).

This can be a hard problem, but is not, if one has both the necessary schemata for various parts of the problem, and appropriate descriptors to help get them selected and interrelated. If, for example, one knows that the dot produce $\vec{u}.\vec{V}$ between any vector $\vec{V}$, and a unit vector $\vec{u}$ in some special direction, gives us the component of $\vec{V}$ in the direction of $\vec{u}$, one can find any vector from O to the plane

(for example, the vector (5,0,0)), and dot it into a unit normal vector. If one has any non-zero normal vector $\vec{N}$, one can get a unit normal vector by dividing $\vec{N}$ by its length. If one has the equation of the plane in the form ax+by+cz=d, one easily gets a normal vector $\vec{N}$. And so on. Given descriptors such as these, the problem of selecting and sequencing the necessary schemata becomes entirely manageable.

f. Frame Adequacy Judgements. When a frame is selected and retrieved, and when a mapping of input data into frame variables has been attempted, there must be a mechanism for deciding either that the "fit" is a satisfactory one, or else that it is not.

2. When we began our studies, we expected that a major difference between stronger and weaker students would be differences in "creative" problem analysis, often based on superior heuristic reasoning. This was based on observations of the problem-solving behavior of adult experts, for whom (for example) the problem of solving the equation

$$3x+10+\frac{7}{x} = 2x+15+\frac{1}{x}$$

might be handled by an analysis along roughly these lines:

I don't like that "x" in the denominator. Can I eliminate this feature? Yes, multiply both sides of the equation by x.

and so on.

Somewhat similar behavior CAN (occasionally) be observed also in pre-college students, but considering the large body of our interviews, one has to conclude that this kind of difference does not explain very much of the difference between strong and weak students. We would add to the list at least seven additional

differences, four of which deal with frames:

i) Strong students possess a large repertoire of frames (although the difference between strong and weak students in terms of possession of frames seems not so great as one might expect);

ii) Strong students recognize the importance of frames. The difference here is more striking. R, an adult studying college algebra at a community college, is a typical "weak" student, having considerable difficulty in learning mathematics, despite high motivation and tenacity. R was confronted by the need to compute

$$\left(\sqrt{x+a} - \sqrt{x-a} \right)^2$$

by carrying out the squaring. When the instructor referred to $(A+B)^2$, R considered this an irrelevant digression, and did not see how to make use of it. This was a typical response by R. Apparently R has not recognized that one has certain key patterns in one's mind - call them semantic templates - and that mathematics sometimes requires an effort to match up some notation on paper with an appropriate semantic template. R feels that something explicitly present in

$$\left(\sqrt{x + a} - \sqrt{x - a}\right)^2$$

should determine the next step. In effect, she expects all of mathematics to resemble the long division algorithm, to be a rote VMS sequence.

A slightly similar, though far less extreme, situation is readily observable among tenth grade students trying to find the equation of an ellipse from the "sum of the distances

from two fixed points" definition; one easily gets, by a step-by-step process, to the result

$$a^2x^2 + c^4 = c^2x^2 + a^2c^2 + c^2y^2. \qquad \text{(i)}$$

Nothing in this line clearly determines what should be done next. To get further, one must shift gears and begin to think of the desired result. If one knows that the goal is an equation of the form

$$\frac{x^2}{R^2} + \frac{y^2}{W^2} = 1, \qquad \text{(ii)}$$

for some R and some W, then getting from equation (i) to the final result is easy - but what is required is a change from step-to-step processing, switching instead to a strategy of keeping the goal clearly in mind.

iii) Strong students are skillful in retrieving the appropriate frames; they see

$$e^{2t} - 5e^t + 6=0$$

as a quadratic equation, they see

$$\sqrt{(x+a)^2 + y^2} + \sqrt{(x-a)^2 + y^2} = 2c$$

as requiring

$$(A+B)^2,$$

and so on.

iv) Strong students have better mechanisms for determining the adequacy of frame selections.

v) Perhaps the most impressive skill of very strong students is their possession of a powerful meta-language, and their skill in using it.

For example, Davis et al. [1978] reports interview protocols that indicate that the most able ninth graders observed (who had studied some calculus intermixed with their

ninth-grade algebra) were able to deal with a complicated VMS sequence not merely as something to do (a "verb"), but also as a thing to discuss and to analyze (a "noun"); this process is aided somewhat if an appropriate meta language is provided (though the most able students can create a meta language without outside help). For example, consider the problem: A ball is thrown straight up so that it reaches a height of 40 feet. How fast is it thrown? The weakest students cannot solve this problem. Somewhat stronger students can solve it, by carrying out an appropriate VMS sequence, following procedures demonstrated in class, but cannot discuss the problem abstractly without actually carrying out each step. The (few) strongest students can discuss and analyze the solution correctly, in detail, without carrying out each step. For example, somewhat paraphrased, here is an actual protocol from a ninth-grade student of high ability:

"From $h''(t) = -32$ ft/sec^2, you can get $h'(t,V_0)$ with an unknown constant V_0. But for h to be a maximum, h' must be zero. The equation $h'(t_{max},V_0)=0$ gives t_{max} in terms of V_0. Integrating $h'(t)$ gives $h(t,V_0)$. Now, for $t=t_{max}$, $h(t_{max},V_0)=40$, and since t_{max} can be expressed in terms of V_0, we can solve for V_0." [This is not verbatim, but does follow the answer of the student, talking accurately, in detail, about how he would solve the problem - without executing the steps of the actual VMS sequence, and hence without benefit of the sequence of visual cues as the

written solution develops. The student's actual language made use of phrases such as "plug in this expression for t-max," and so on.]

vi) Strong students recognize the importance of careful, painstaking work. [This may seem obvious, but it's important not to get so caught up in the consideration of "cognitive" differences that one loses sight of certain very important "non-cognitive" matters.]

vii) Strong students have a different view of the epistemology of mathematics. Examples are given in Davis, Jockusch, and McKnight [1978], and Erlwanger [1973].

3. David Page, some years ago, suggested that a main goal of effective mathematics teaching is to get most students to deal with mathematics the way the best students already do. To what extent is the data processing of strong students "teachable" to other students? Our work thusfar has dealt only peripherally with questions of teaching, but some of these peripheral observations are encouraging. Teachers working with the project do find themselves using project observations as a basis for revising their teaching procedures - for example, they now work harder to teach students a "meta" language. Consider the following example, reported in McKnight [1979], which is in part a teaching use of Erlwanger's data on super-procedure errors:

We have explored procedures, especially the relation of super-procedure to sub-procedures, in looking at the learning of some of the more algorithmic aspects of school algebra. One technique in those explorations might be called the "before and after snapshot." Brief (about five minutes) quasi-clinical interviewing was used to generate talk-aloud protocols for students solving two or three routine problems and to gather student

clarifications of ambiguous points in the protocol. "Before" and "after" pictures were obtained by conducting two interviews on the same type of problem, the first within the first day or so of instruction relevant to the problem type and the second towards the end of immediately relevant instruction. One such "before and after snapshot" is discussed here, but we have collected many such protocols, of which this one is fairly typical.

The subject was a seventh-grade boy, whom we shall call "P", who had been in an introductory algebra course for about two and a half months. The course had covered signed numbers, some beginning analytic geometry, logic and formal derivations of algebraic expressions and it was just then moving into instruction on simplifying algebraic expressions prior to going into the arithmetic of polynomials. Thus, tasks like

"Simplify $4x^2 + 2x - [x^2 + 3x]$"

or

"Simplify $5x^2 + 3x - [x^2 + 4x]$"

represented new, challenging problems although some preparation had been made during the course of instruction on algebraic derivations (which involved use of formal axioms like the distributive law, commutative law of addition, etc.)

Let us first look at P's "after" picture which came on the eighth day of instruction on the topic just before moving to the next topic. Given the task, "Simplify $5x^2 + 3x - [x^2+4x]$," his protocol may be summarized as follows:

(1) He first announced that he was going "to make the minus into plus a negative" and wrote

$$5x^2 + 3x + -[x^2 + 4x].$$

[Notice, he unhesitatingly calls up and executes this simple syntactic transformation procedure. He possesed a verbal

label for what he did, had no doubts that it was a legal syntactical move, and knew that it would be helpful to him overall, suggesting that he has a clear super-procedure embedding this step.]

(2) Next he announces that he is going to handle the "parenthesis" and writes

$$5x^2 + 3x + -x^2 + -4x.$$

[Again, a clearly legal syntactic transformation based on the distributive law which P executes without question or hesitation, indicating the functioning of an organized super-procedure. Instruction had suggested this transformation as one of several that would be legal and effective in what, in meta-language, would be called "removing parentheses." Stress was placed on a meta-language discussion of the super-procedure rather than extensive drill on executing one form of it. Notice P calls the brackets here "parentheses" even though the instructor and the interviewer had called them "brackets." The meta-language discussion has "taken," at least in part.]

(3) P next uses the commutative law of addition twice to bring his "like" terms together, writing

$$5x^2 + {}^{-}x^2 + 3x + {}^{-}4x.$$

[Again, he does this unhesitatingly and without pause between the two executions indicating a clear super-procedure. He mentions that he can now do this since it's "addition," which suggests he has some awareness of why the first two steps were necessary. The effect of formal derivations on P's procedure is interesting. Further work typically curtails a step like this one, leaving it implicit in the step which follows, "combining like terms." Thus, this step indicates that the structure of P's super-procedure has been shaped by his experience in formal algebraic derivatives in which he has made moves like this numerous times. This is strange in comparison to (1) the fact that, asked to name the justifying law here, he comes up finally with the associative law rather than the commutative and (2) the confusions in this "before" protocol which will be discussed below.]

(4) Finally, he twice adds "like" terms, writing the final result,

$$4x^2 + {}^{-}1x.$$

[Again, clear super-procedure and correct sub-procedures. Interestingly, the final "$^{-}1x$" in his answer reveals the carefulness and precision of his syntactical transformations. Such carefulness is also curtailed with further experience and confidence. Again, such clarity about syntactical matters will be seen to be in marked contrast to his "before" protocol.]

This "after" protocol seems to be obvious and to be a simple exercise in rhetoric application. The usefulness of the approach becomes clear in contrasting the "before" protocol to this precise, efficient, correct target performance in the "after."

The "before" protocol was taken at the beginning of the third (45 minutes) period of instruction on the topic. P had worked simpler exercises and had seen worked exercises something like his assigned task, "Simplify $4x^2 + 2x - [x^2 + 3x]$."

His "before" protocol may be summarized as follows:

(1) P announces that he wants to do something with the "like terms," identifying $4x^2$ and x^2 as such, but he is concerned that "it's minus that [x^2] so...?" Later he expresses confusion about the fact that the 4 and the x^2 in the $4x^2$ are combined by multiplication while the 3x and the x^2 in $x^2 + 3x$ are combined by addition. [Notice, P is at first unable to call any sub-procedure; his "so..." trails off into a question. He has no super-procedure ready at hand. At best, he seems to have some heuristic guidance in the form of knowing he needs to "combine like terms." This leads him to reflect on the expression and his reflecting is a strange mixture of syntactic awareness and confusion. He realizes the minus should tell him something but he's not sure what. He contrasts the connectives in $4x^2$ and $x^2 + 3x$, which is astute in the sense that most similar students would not have that clear a grasp of the syntax of $4x^2$ or be led to compare the two sub-expressions. Yet, the contrast is almost esoteric in its irrelevance to the needed procedure and shows a confusingly wide-ranging, undirected reflection.]

(2) After noting "you can't add," he copies down $4x^2 + 2x$ without combining unlike terms as he briefly considered, and then goes on to complete the expression,

$$4x^2 + 2x - 3x + x^2.$$

[P's explicit heuristic of combining "like terms" (apparently with a further implicit heuristic of bringing "like terms" in proximity as we see correctly done in a full-blown form in step 3 of the "after") and his syntactical confusions lead him to take the plunge with an incorrect, ad hoc procedure. He has the syntax of $4x^2 + 2x$ correct and even the commutation of $x^2 + 3x$ would be correct inside parentheses. He seems to have no awareness of the syntactical function of the parentheses nor any legitimate procedure for removing them. Thus the ad hoc move that is guided only by rough heuristics, not an organized super-procedure, and which shows lack of knowledge of a key syntactical point or else a lack of use of such knowledge in a validity check of the simple ad hoc procedure he has concocted (to relieve the pressure on him?).]

(3) P next said that he had to "take" the 3x from both the 2x and the $4x^2$. He then wrote a $^-1$ below the 2x and indicated that 2x - 3x gave $^-1$. When questioned further, he stated that "you minus the x [in 3x] from 2x." [Having now followed his rough heuristics to the point where they can now dictate directly a step, P proceeds to follow up his syntactical blunder in the preceding step with a further syntactical blunder. This is a clear case of some sort of regression under strain because in previous days P had easily done related simpler problems correctly using the same procedure which would be the appropriate sub-procedure here. Thus, encountering sub-problems which he should be able to do in a context that is already demanding much of his cognitive resources causes malfunctions in his syntactical understandings and in his sub-procedures.]

(4) P now moves confidently to take 3x from $4x^2$ writing 1x below the $4x^2$. Then he suddenly expresses some doubt that this is correct and, after a pause, decides it is not, because "$x^2 - x$" doesn't "equal x." He then goes back to consider whether he needs to take the 3x from both terms and eventually decides he does not. [The confidence with which he begins this step seems to show that P has at least a rudimentary super-procedure to cover a few steps. The lack of hesitation seems a clue to such super-procedures. Suddenly his rough heuristic about "like terms" seems to enter in a different way, acting as a validity check which leads him to reject what he has just done. It is not clear why he is able to go on to realize that he was making a previous error in even taking 3x from $4x^2$. A reasonable hypothesis, based on previous evidence of the good level of knowledge P has about algebraic syntax, is that it is this syntactical knowledge that surfaces here. We have often seen that one error will start children on a chain of errors like some short-term, localized algebraic dementia. P's earlier wrong transformation seems to have done this, but encounter with some syntactical fact of which he is sure brings him out of it and leads him to correct another of his errors.]

(5) P then rewrites $4x^2 + 2x + 3x + x^2$ to give

$$4x^2 - x^2 + 2x - 3x.$$

[Somehow, "like terms" has put him back on the right track and he can access and use the same sub-procedure that he uses in step 3 of the "after" and, if the previous steps had been correct, it would be appropriate here. Unlike step 3 of "after," this sub-procedure is not executed as a part of a super-procedure but rather as a part of something like a "visually moderated sequence" (Davis,1979) in which visual and other outputs from a previous step lead to selection and execution of a sub-procedure even without an overall plan.]

(6) Finally, P moves quickly to write 4 as the output of $4x^2 - x^2$ and $^-1$ as the output of 2x - 3x, giving him a final answer of $4 + ^-1$, which does not trouble him at all. [This, of course, is the application twice of the same inappropriate procedure seen in step 3 of this protocol. It is mysterious why P can snap out of pressure- or error-induced regression enough to correct two errors but not go on to correct a third. It certainly goes to show the persistence of some errors. One possible difference is that while P has performed a sort of validity check both in step 3 and in step 5, that of step 5, "$x^2 - x$" is correct and shows his

error while that of 3, "you must minus the x from the 2x" is a syntax-ignoring, misleading verbalization. Such a verbalization may have short-circuited the validity check process and left the error-producing sub-procedure intact.]

While P's final answer is, of course, atrocious and while he showed several shortcomings in this "before" performance, his actual errors are few and easily corrected. It simply took pointing out the error in steps 3 and 6 to make P realize that it was wrong and, even to him, silly [although other students have proved more resistant to such knowledge], and a brief bit of instruction about the syntax of parentheses and about their removal. At least one practical benefit of the super/sub-procedure analysis is to pinpoint errors sufficiently precisely to greatly aid in remediation.

Comparison of P's "before" and "after" protocols suggests several conclusions (which are supported by other related protocols) including the following:

(1) The key task in tasks like that of P is really that of <u>super-procedure formation</u>. Probably the strongest contrast between before and after is that of the confidence and lack of hesitation in the after which is justified by the correct sequence of sub-procedure accessing.

(2) Execution of already formed procedures as sub-procedures in such a context (i.e., super-procedure formation) often results in a <u>decay</u> in performance of the already-formed procedures. This is similar to what Bruner and others have found in motor skills learning. It may make a case for over-learning of procedures that will be sub-procedures in a variety of more complex procedures.

(3) The first two phenomena combine in a change which might be called <u>mutual clarification</u>, i.e., when a correct super-procedure is stably formed, so will be correct sub-procedures. While it was not wholly the case with P, in some cases we have seen group instruction which continued to focus on forming the correct super-procedure result in seemingly spontaneous correction of sub-procedures.

(4) Formation of super-procedures seems clearly to run from trial-and-error crystallization around a few heuristic guides to a stable, stereotypic structure. This might be called an exodigmatic (paradigm from "within") formation approach as contrasted to an hypodigmatic (paradigm from "above") approach (i.e., repeated imitation of an overall model).

(5) The syntax of algebraic expressions may be a key or milestone kind of knowledge in algebraic learning. The degree of syntactic security seems to be a crucial element in a student's predisposition to regression under strain.

(6) Formal derivations (such as these students had been exposed to) may help in creating secure syntactic awareness and transformations even though it seems often to fail at the task of making the student able to provide explicit justifications.

For further examples, please refer to Davis, Jockusch, and McKnight, 1978, to Davis [1979-A], and Davis [1979-B].

REFERENCES

Binet, A. L'étude expérimental de l'Intelligence. Paris: Schleicher Freres, 1903.

Bobrow, D. "Natural Language Input for a Computer Problem-Solving System," in: M. Minsky, ed., Semantic Information Processing. Cambridge, MA: MIT Press, 1968.

Brown, John Seely and Richard R. Burton. Diagnostic Models for Procedural Bugs in Basic Mathematical Skills, in Cognitive Science, Vol. 2, No.2. April-June, 1978.

Davis, Robert B., Elizabeth Jockusch, and Curtis McKnight. "Cognitive Processes in Learning Algebra" in The Journal of Children's Mathematical Behavior, Vol. 2, No. 1 (1978).

Davis, Robert B. "Representing Knowledge About Mathematics for Computer-Aided Teaching, Part I. -Educational Applications of Conceptualizations from Artificial Intelligence," in E.W. Elcock and Donald Michie, eds, Machine Intelligence 8: Machine Representations of Knowledge. Ellis Horwood, Ltd. (dist. by Halstead Press, John Wiley & Sons, Inc., New York), 1977, pp. 363-386.

Davis, Robert B., "Conceptualizing the Structures Underlying Cognitive Behavior - The Usefulness of "Frames". Paper presented at A.E.R.A. Meeting, San Francisco, Spring, 1979B.

Davis, Robert B., "Error Analysis in High School Mathematics, Conceived as— Information-Processing Pathology." Paper presented at A.E.R.A. Meeting, San Francisco, Spring, 1979A.

Erlwanger, Stanley H. "Benny's Conception of Rules and Answers in IPI Mathematics" The Journal of Children's Mathematical Behavior, Vol. 1, No. 2 (Autumn, 1973), pp. 7-26.

Erlwanger, Stanley H. Case Studies of Children's Conceptions of Mathematics. Doctoral Dissertation, University of Illinois, Urbana, Illinois, 1974.

Erlwanger, Stanley H. "Case Studies of Children's Conceptions of Mathematics- Part I, "The Journal of Children's Mathematical Behavior, Vol. I, No. 3 (Summer, 1975) pp. 157-283.

Hinsley, Dan, John Hayes, and Herbert Simon. "From Words to Equations - Meaning and Representation in Algebra Word Problems," in: Marcel Just and Patricia Carpenter, eds., Cognitive Processes in Comprehension. Hillsdale, N.J.: Erlbaum, 1977.

Kilpatrick, Jeremy. Analysing the Solution of Word Problems in Mathematics: An Exploratory Study. Ph.D. thesis, Stanford University, 1967.

Kolata, Gina Bari "Blood Sugar and the Complications of Diabetes," Science, Vol. 203 (March 16, 1979) pp. 1098-9.

McKnight, Curtis. "The Use of "Frames" in Conceptualizing Learning and Problem Solving in Mathematics." Paper presented at A.E.R.A. Meeting, San Francisco, Spring, 1979.

Michener, Edwina R. Understanding Understanding Mathematics, Cognitive Science, Vol. 2, No. 4 (Oct.-Dec., 1978), pp. 361-383.

Minsky, Marvin. "A Framework for Representing Knowledge," in: P. Winston, ed., The Psychology of Computer Vision. New York: McGraw-Hill, 1975.

Minsky, Marvin, and Papert, Seymour, Artificial Intelligence Memo No. 252, Cambridge, Mass.: Massachusetts Institute of Technology Artificial Intelligence Laboratory, January 1, 1972.

Paige, Jeffery M. and Herbert A. Simon. "Cognitive Processes in Solving Algebra Word Problems," in:(B. Kleinmuntz, ed., Problem-Solving: Research, Method, and Theory.) New York: John Wiley, 1966. pp. 51-119.

Rumelhart, David E. and Andrew Ortony, "The Representation of Knowledge in Memory," chapter 4 in: Richard C. Anderson, Rand Spiro and William Montaque, eds., Schooling and the Acquisition of Knowledge. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Schank, Roger. "Representation and Understanding of Text," chapt. 27 in: E.W. Elcock & Donald Michie, eds., Machine Intelligence 8: Machine Representations of Knowledge. Ellis Horwood Ltd. (dist. by Halstead Press, John Wiley & Sons Inc., New York), 1977, pp. 575-619.

Van Lehn, Kurt and John Seely Brown. Planning Nets: A Representation for Formalizing Analogies and Semantic Models of Procedural Skills. (to appear).

Winograd, Terry. Procedures as a Representation for Data in a Computer Program for Understanding Natural Language. Cambridge, Massachusetts: M.I.T. Artificial Intelligence Laboratory, Feb. 1971.